

国立国語研究所学術情報リポジトリ

# 『東京語アクセント資料』と辞書アクセント：尾高型アクセントを事例とした資料評価

著者	相澤 正夫
雑誌名	日本語科学
巻	1
ページ	80-91
発行年	1997-05
URL	<a href="http://doi.org/10.15084/00001968">http://doi.org/10.15084/00001968</a>

# 『東京語アクセント資料』と辞書アクセント

## —尾高型アクセントを事例とした資料評価—

相澤 正夫

(国立国語研究所)

### キーワード

東京語アクセント, 尾高型アクセント, 辞書アクセント  
資料評価, 『東京語アクセント資料 上・下』

### 要 旨

『東京語アクセント資料』をアクセント研究に有効に活用するためには、まず第一に、目的に応じた資料評価を十分におこなう必要がある。本稿では、資料評価のあり方とその方法を模索するために、事例として東京語の尾高型アクセントを取り上げ、『東京語アクセント資料』におけるその出現状況を問題とする。具体的には、既刊の4種の辞書のアクセント情報と対照させながら、それらとの異同を詳細に調査し、アクセントの計量的研究における資料面での問題点を指摘する。

### 1. はじめに

『東京語アクセント資料 上・下』(1985) (以下『東京ア』と略称) の主な特徴として、次の4点が考えられる。

- ① 現代東京語でアクセントのゆれが予想される12,803語を収録。
- ② 年齢差, 地域差 (山の手と下町), 男女差に配慮して選定した19名の話者について, 個人別に情報を記載。話者はアルファベットで識別。大文字は男性, 小文字は女性, 「山」は山の手, 「下」は下町, 数字は生年 (西暦) の下2桁。例えば, A氏は男性・下町出身・1962年生まれ。以下同様。

A 下62, b 山59, C 下58, d 山58, E 下53, F 下50, g 下47, H 山43, i 下43,  
J 山39, K 下39, l 山35, m 山35, N 山30, o 山30, P 下29, q 山29, r 山20,  
s 山11

- ③ 既刊の辞書4種に記載されているアクセント型と対照。辞書は以下のように略称で表示。

『新明解』: 『新明解国語辞典 第3版』(三省堂)

『NHK』: 『日本語発音アクセント辞典』(日本放送出版協会)

『明解ア』: 『明解日本語アクセント辞典 第2版』(三省堂)

『全国ア』: 『全国アクセント辞典』(東京堂出版)

- ④ 準備段階でアクセントをチェックした2名のアクセントを併載。2名の属性は次のとお

り。

X山55, y下20

相澤 (1996a, 1996b) では、①②を「尾高型アクセントの衰退動向の多角的分析」というテーマに有利と考え、『東京ア』をデータ収集の対象として利用した。①によって語数の多さ、したがって語ごとの特徴（語構成、品詞、語種など）の多様性が、また、②によって話者の属性の多様性が確保されるからである。一方③④は、有用ではあるが、副次的な情報に止まるものと考えた。例えば、結果として③からは辞書それぞれの個性（アクセントのゆれに対して記述的か規範的かなど）を知ることができるし、④からは試掘的な調査が妥当であったかどうか、調査法を評価する際の手がかりが得られるものと予想したが、テーマとは区別して扱った。

データ採集に際しては、次の3点の条件を全て満たすこととした。

- ① 4種の辞書が、全て尾高型アクセントを登録していること。
- ② 語形が一定していること。
- ③ 19名分の情報が全て揃っていること。

ここで、②は語音の違いによる影響を排除するための条件、③は計量的・統計的な処理をするために必要な条件である。これに対し、①はいわば作業上の便宜的な措置であった。その語が確かに尾高型で発音されていたことを保証するために設定した、任意の一条件にすぎないとも言えるからである。しかし、『東京ア』によって尾高型アクセントの衰退動向を端的に捉えるための方法としては、きわめて有効であろうと予測した。

結果として上述の目標はほぼ達成されたが、相澤 (1996a) の「おわりに」で指摘したように、いくつかの課題が残された。その一つが、「4種の辞書が揃って尾高型アクセントを記載することを条件に語例を採集したが、それが妥当であったかどうか検証する必要がある。辞書における尾高型の記載について、①3種が記載する語、②2種が記載する語、③1種が記載する語、④4種の辞書は尾高型を記載しないが、19人の話者に尾高型が現われる語、以上四つの観点から再度語例を採集し、今回の結果と突き合わせながら異同を評価する必要がある。」というものであった。これは、分析対象としたデータの制約を知るという意味で、なおざりにできない課題である。

本稿では、この課題に応えるために、まず、尾高型アクセントの出現状況をくまなく調べ上げ、その実態を整理して示す。次に、『東京ア』と辞書アクセントとを対照させることによって、アクセントの計量的研究における若干の問題点を指摘し、資料評価の必要性を論ずる。

## 2. データの収集

ここでは、相澤 (1996a) での調査を「前回調査」、本稿のための調査を「今回調査」と呼ぶ。前回調査は局所調査であり、今回調査は全体調査である。なお、今回調査では、尾高型アクセントの出現状況のみを問題とし、併存するアクセント型との関係は扱わないことにする。

今回調査では、『東京ア』における尾高型の出現状況を次のような方法で調査し、該当語を採集した。語形が一定していることと、19名分の情報が全て揃っていることは、当然の前提条件である。<sup>1</sup>

① 4種の辞書のうち1種でも尾高型を記載していれば、その語を採集する。

② 19人の話者のうち1人でも尾高型で発音していれば、その語を採集する。

ここで、①の該当語グループを「辞書アクセントグループ」（以下《辞書ア》と略記）、②のそれを「19人アクセントグループ」（以下《19人ア》と略記）と呼ぶことにすれば、該当語の全体は《辞書ア》と《19人ア》の和集合となる。前回調査では、「4種の辞書が全て尾高型を登録していること」を条件として語例を採集したが、そのときの該当語は《辞書ア》に、部分集合として含まれることになる。

結果として採集された該当語は、3拍語796例、4拍語1,337例、5拍語475例の、総計2,608語であった。これは、『東京ア』の全収録語（12,803語）の20.4%、約5分の1にあたる。このうち839語は前回調査で分析対象とした語であるから、今回調査で新たに1,769語が追加されたことになる。このように、最もゆるい条件で採集することによって、尾高型アクセントの関与するとみられる語が、おそらく最大限に採集されたといっておよさう。<sup>2</sup>

以上、グループ間の関係と該当語数をまとめて示せば、図1のようになる。

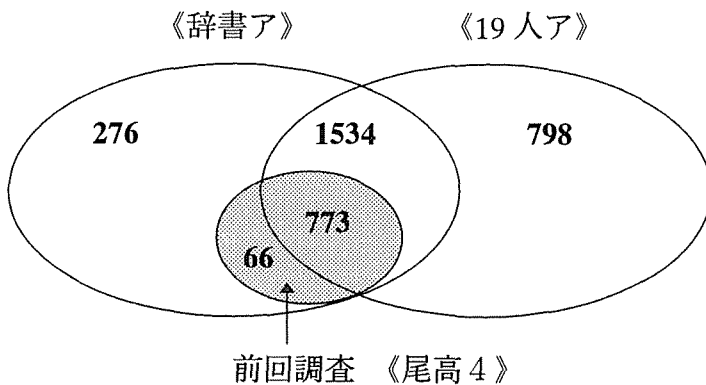


図1 分析対象とするデータ

### 3. 結果とその分析

#### 3.1. 全体の概観

まず、今回調査の結果得られた総計2,608語について、19人の話者全体の尾高型出現率を求めてみる。計算式は次の通り。

尾高型出現率＝尾高型の全出現度数／全体の度数

ここで、「尾高型の全出現度数」とは、語ごとの尾高型の出現度数（0～19度）を総計2,608語分すべて合算した数値、「全体の度数」とは、尾高型の出現機会の最大値（19人×2,608語）をさす。

結果は、21.0%（10,417／49,552）となった。これを全19人のうちの何人という形で示せば、19人中の4人ということになる。前回調査の結果は、尾高型出現率が29.9%、19人中の6人弱程度であったから、今回調査で新たに1,769語を追加したことによって、尾高型出現率はかなり低下したことが分かる。

さて、今回調査の《全体》2,608語は、4種の辞書における尾高型の記載状況の違いによって、次の5グループに分けることができる。括弧内に該当語数を示す。

《尾高4》：4種の辞書全てが尾高型を記載するグループ（839語）

《尾高3》：4種の辞書のうち3種が尾高型を記載するグループ（434語）

《尾高2》：4種の辞書のうち2種が尾高型を記載するグループ（213語）

《尾高1》：4種の辞書のうち1種が尾高型を記載するグループ（324語）

《尾高0》：4種の辞書全てが尾高型を記載しないグループ（798語）

《尾高4》は前回調査で分析対象としたグループそのものである。そこで、それ以外の4グループについても尾高型出現率を求めてみよう。計算式は、基本的に前述のものと同じである。《全体》2,608語についての数値を、グループごとの尾高型の全出現度数とその該当語数に置き換えて、それぞれ計算すればよい。すなわち、任意のグループ《尾高X》の尾高型出現率を、次の計算式で求める。

《尾高X》の尾高型出現率＝《尾高X》の尾高型の全出現度数／《尾高X》の全体の度数  
結果は、次の通りである。ほぼ5%の間隔で、段階的に減少していくことが分かる。

《尾高4》    《尾高3》    《尾高2》    《尾高1》    《尾高0》

29.9%   >   25.6%   >   20.5%   >   15.5%   >   11.6%

これは、4種の辞書において尾高型を記載するものが多いほど、『東京ア』の19人話者における尾高型の出現率が高いことを示しており、両者には正の相関関係が認められる。したがって、前回調査で尾高型の衰退動向を観察するために、あえて《尾高4》に絞って局所調査をおこなったことは、決して的外れな選択ではなかったことが分かる。結果的に、あまり衰退の進んでいない語をも最大限に分析対象とすることができたという点では、むしろ変異の中に変化の過程をつぶさに観察するという目的に適合していたとみるべきかもしれない。

辞書の1票は、あくまでも話者個人の1票とは異質なものである。しかし、そのことを弁えたうえで方法さえ誤らなければ、辞書をあたかも個人のようにみなして、アクセントの計量的な分析に援用することはできそうである。これは、そんなことを示唆する一事例ではあるまいか。

### 3.2. 拍数別の比較

ここでは、語の長さ（＝拍数）の違いによって、尾高型の出現状況に差がないかを検討する。《全体》2,608語を、《3拍》796語、《4拍》1,337語、《5拍》475語の3グループに分け、3.1.で用いた、辞書の記載状況によるグループ分けにしたがって、尾高型出現率（%）を比較すると、次ページ上段のようになる。括弧内に該当語数を示す。また、図2は、それをグラフ化したもので、《全体》の出現率も一緒に示した。

図2から明らかなように、《3拍》《4拍》《5拍》をまとめた《全体》は、3.1.で確認したとおりに、右下がりのきれいな直線を描いている。これを基準にして見ると、概略、《3拍》はそれを上回りながら、また《4拍》はそれを下回りながら、やはり右下がりに近い線を描いている。これに対して、《5拍》は《尾高3》だけが異様に高く、その他はむしろ横這いに近い状態である。

	《全体》	：	《尾高 4》	《尾高 3》	《尾高 2》	《尾高 1》	《尾高 0》
《3 拍》	29.3	：	43.8	31.9	23.5	23.1	12.6
	(796)		(306)	(127)	(45)	(74)	(244)
《4 拍》	16.6	：	22.8	18.7	20.3	12.4	9.9
	(1,337)		(431)	(207)	(115)	(166)	(418)
《5 拍》	19.6	：	18.6	31.7	18.5	14.8	14.8
	(475)		(102)	(100)	(53)	(84)	(136)
《全体》	21.0	：	29.9	25.6	20.5	15.5	11.6
	(2,608)		(839)	(434)	(213)	(324)	(798)

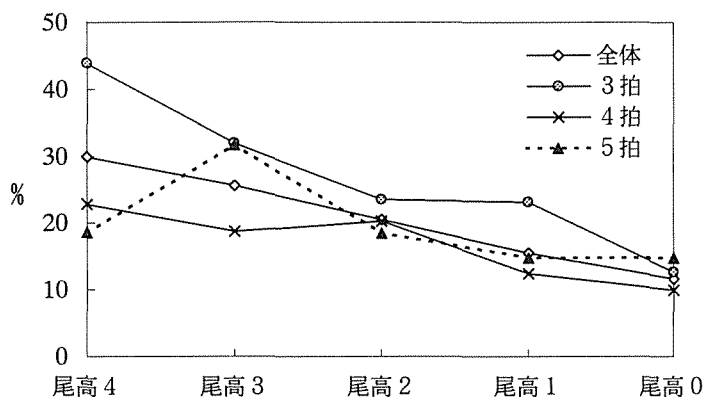


図 2 拍数別の尾高型出現率の比較（全体）

前回調査では《尾高 4》のみを対象として、3 拍、4 拍、5 拍の順に尾高型の出現率が低下する（但し、4 拍と 5 拍の差は微妙）という一応の結論を得たが、今回調査の結果によれば、新たに追加された《尾高 3》《尾高 2》《尾高 1》《尾高 0》は、いずれもこのような序列に従っているとは言いがたい。また、今回調査の《尾高 4》《尾高 3》《尾高 2》《尾高 1》《尾高 0》をまとめた《全体》の尾高型出現率は、《3 拍》29.3%、《4 拍》16.6%、《5 拍》19.6%となつて、むしろ《4 拍》と《5 拍》がわずかに逆転しており、ここにも前回調査のような序列は見出せない。

そうだとすれば、前回調査の結論は、あくまでも《尾高 4》を対象とした局所調査の結果にすぎないことになるのであろうか。

#### 4. 資料評価

##### 4.1. 辞書にのる語・のらない語に関わる問題

今回調査では、《辞書 A》の語例採集に際して、「4 種の辞書のうち 1 種でも尾高型を記載していれば、その語を採集する」という方針をたてた。そこでおこなったのは、実は「尾高型の記載のある語の選出」であつて、尾高型の記載のない場合の事情については全く立ち入っていない。しかし、記載がないといっても、実際には次のような二つの場合があることに注意しなければな

らない。

① 4種の辞書全てにその語がのっているが、一部あるいは全部に尾高型の記載がない場合。

② 4種の辞書の一部あるいは全部に、そもそもその語がのっていない場合。

ここでは、①の該当語グループを《欠なし》、②のそれを《欠あり》と呼ぶことにする。《欠なし》の場合は、特定の辞書がその語について尾高型の記載を積極的に排除しているとみなすことができるが、《欠あり》の場合は、そもそもその語が特定の辞書に立項されていないのであるから、尾高型の状況について云々することはできないはずである。

そこで、《欠あり》は尾高型について不明な部分を含むので集計から除外し、《欠なし》について、辞書の記載状況によるグループ分けにしたがって、尾高型出現率を比較してみると、次のようになる。括弧内に該当語数を示す。図3は、それをグラフ化したものである。

	《全体》	《尾高4》	《尾高3》	《尾高2》	《尾高1》	《尾高0》
《3拍》	30.2 (658)	43.8 (306)	30.3 (80)	22.5 (25)	19.7 (52)	12.5 (195)
《4拍》	16.6 (991)	22.8 (431)	17.0 (126)	15.3 (59)	10.1 (103)	9.2 (272)
《5拍》	14.6 (292)	18.6 (102)	24.7 (54)	8.6 (28)	2.6 (38)	9.9 (70)
《全体》	20.9 (1,941)	29.9 (839)	22.7 (260)	15.2 (112)	11.2 (193)	10.5 (537)

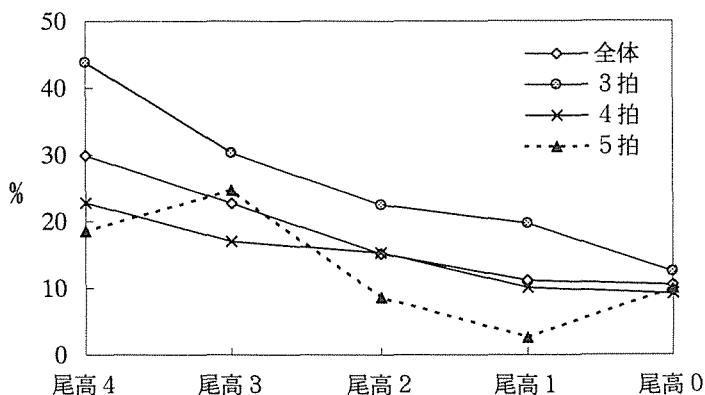


図3 拍数別の尾高型出現率の比較 (《欠なし》のみ)

図3から明らかなように、《欠なし》の場合も、《3拍》《4拍》《5拍》をまとめた《全体》は、ほぼ右下がりの直線を描いている。これを基準としてみると、概略、《3拍》はそれを大きく上回りながら急な右下がりの直線を描いており、《4拍》は《尾高4》《尾高3》では《全体》を下回るが、やがて重なるようにしながら右下がりの緩やかな線を描いていることが分かる。これに対

して、《5拍》は《尾高3》だけがやはり異様に高い出現率を示し、その他の部分でも凹凸が目立っている。

《欠なし》の場合に注目されるのは、《尾高4》《尾高3》《尾高2》《尾高1》《尾高0》をまとめた《全体》の尾高型出現率が、《3拍》30.2%、《4拍》16.6%、《5拍》14.6%と、前回調査で観察されたような語の長さによる序列を示していることであろう（但し、4拍と5拍の差はやはり微妙）。図3からも分かるように、これには《尾高2》《尾高1》において、はっきりとした序列が現れていることが大きく影響しているものと思われる。

以上のように、分析対象から《欠あり》を除外することによって、辞書のアクセント情報から曖昧な部分を排除することができる。辞書アクセントと19人話者のアクセントの関係をより正確に捉えることを目指すのであれば、このような手続きはむしろ不可欠と言えよう。その意味で、前節における《全体》2,608語の分析はやや大雑把であり、資料の内容に対する評価が不十分であったと言わざるをえない。

#### 4.2. 複合語に関わる問題

図2、図3ですで見たとように、《5拍》の特に《尾高3》は、異様に高い尾高型出現率（図2では31.7%、図3では24.7%）を示している。これは、一体何に由来するのであろうか。

一つには、ある一定の特徴をもった語群が出現率を引き上げているのではないかと予想される。そこで、《5拍》の《尾高3》に属する100語を実際に調べてみると、所属語の特徴に大きな偏りがあることが判明した。すなわち、同じ後部成素をもつ複合語が多数含まれていたのである。具体的には、「～・書」「～・所」「～・署」を後部成素とする複合語だけで26語に及び、この語群における尾高型出現率は、67.2%にも達している。これに対して、残りの74語における出現率は19.3%にすぎないから、特定の複合語によっていかに出現率が高められているかが分かる。<sup>3</sup>

次に、該当語例を掲げる。[ ]内の数字は、19人中の何人が尾高型であるかを示す。

「～・書」：[16] 借用書，[15] 証明書，新刊書，申請書，[14] 引用書，[13] 参考書，内申書，入門書，領収書，[12] 明細書，[11] 説明書，福音書，[10] 受取書，見積書（以上14語）

「～・所」：[13] 研究所，講習所，宿泊所，出張所，停留所，養成所，[12] 造船所，[11] 観測所，[10] 休憩所，収容所（以上10語）

「～・署」：[15] 警察署，消防署（以上2語）

このように、生産力の高い（と予想される）後部成素は、それをもつ複合語がどれだけ分析対象に含まれているかによって、結果としての数値に大きく影響することが予測される。『東京ア』は、現代東京語アクセントの実態を知るための調査資料であるから、このような複合語を多く含んでいて当然であるが、辞書の場合は必ずしもそうではない。使用頻度の比較的高い日常語は別としても、一般に複合語アクセント規則で説明のつくものは、むしろ立項しないで済ませるのが普通だからである。したがって、今回調査で追加された語のうち《欠あり》に属するものには、このような生産力の高い後部成素をもつ複合語が、数多く含まれている可能性が高いとみられるので



ある。<sup>4</sup>

表1 「～書(シヨ)」を後部成素とする複合語の尾高型出現状況

表記 : 拍数 : 新明 : NH : 明解 : 全国 : 辞書 : 欠 : X山 : y下 : A下 : b山 : C下 : d山 : E下 : F下 : g下 : H山 : i下 : J山 : K下 : l山 : m山 : N山 : o山 : P下 : q山 : r山 : s山 : 尾高 :

[illegible]

アクセントの計量的分析において、複合語をいかにカウントするかは難しい問題を含んでいる。ここでは、「～・書」を事例として、その一端を垣間見ることにしたい。

表1は、今回調査の結果から「～・書(ショ)」となる複合語を全て抽出し(計52語)、19人話者のうち尾高型で発音する人の合計が多い順に配列したものである。表1で、4種の辞書の欄の「x」は立項なしの意であり、その合計を「欠」に示す。また、「5, 4, 3」の数値は尾高型の記載ありの意、「空欄」は記載なしの意である。「辞書」には尾高型の記載のある辞書の合計を示す。19人話者の欄における数値と空欄の意味は、4種の辞書の欄と同じである。右端の「尾高」には、19人話者のうち尾高型で発音する人の合計を示す。

計52語の内訳をみると、3拍語9例、4拍語8例、5拍語35例と、5拍語が圧倒的に多い。このうち前回調査の対象となっていたのは「声明書、請求書、診断書」のわずか3語にすぎないから、「～・書(ショ)」という複合語は、今回調査で49語も大幅に追加されたことになる。分析対象の自身にこれだけ大きな変化が生ずれば、調査結果に影響が現われるのは当然であろう。

52語全体の尾高型出現率は57.5%、そのうち《欠なし》(24語)のそれは49.1%、《欠あり》(28語)のそれは64.7%である。やはり、辞書にのらない語における出現率の高さが目立つ。このことは、「～・書(ショ)」という後部成素の生産力の高さと決して無縁ではあるまい。

表1からも明らかなように、出現率の高い語(1~44)と低い語(45~52)との間には断絶が認められる。おそらく、前者の前部成素が2形態素からなる語は、複合語アクセント規則の適用によるグループで、後者の1形態素からなる語は、個別の伝承アクセントによるグループと推定される。このことは、前者において、19人話者が大きく2分されていることから支持される。ほぼ一貫して尾高型である人(A, b, d, E, F, g, H, i, m, N, o, sの各氏)と、ほぼ一貫して平板型である人(P, q, rの各氏)とにきれいに分かれ、両者の中間的な人(C, J, K, lの各氏)も存在するからである。

このように、生産力の高い後部成素の場合、複合語アクセント規則の適用によって、極端に言えば無限に近い同型アクセント語群を追加できるのであるから、集計の際にそれらをどう処理すべきかについては、十分な配慮が必要となるのである。

## 5. おわりに

本稿では、尾高型アクセントを事例として『東京ア』の資料評価を試みた。そもそも資料評価とは何か。それに正面から答えるのは難しい。ここでは、「分析の目的に見合った諸条件を、対象とする資料が満たしているかどうかの吟味」ほどの理解で出発した。目的によって手段は異なっている。資料評価の内容も目的によって柔軟に変化しうるものと、基本的には考える。

一方、評価には一定の基準が必要である。しかし、『東京ア』と辞書アクセントの外部に、第三の動かぬ評価基準があるわけではない。現状では両者による相互評価しかありえない。相澤(1996a, 1996b)と同様、本稿では、辞書アクセントを基準に『東京ア』を評価するという方向に偏りすぎたきらいがある。今後は、『東京ア』を基準にして辞書アクセントを評価するという、反対方向の調査研究も必要であろう。

# 注

- 1 相澤 (1996a) と同様、該当語には名詞のほか形容動詞語幹の類も含まれている。
- 2 相澤 (1996a) では、総計832語を対象としている。今回の調査で若干の出入りがあり、総計839語となった。本稿では、修正後の新しい数値を採用する。
- 3 ここで取り上げた《5拍》の《尾高3》に属する100語は、《全体》2,608語から抽出したものである。ちなみに、《5拍》の《尾高X》における①「～・書」「～・所」「～・署」グループだけの尾高型出現率(%)、②それを除外した残りの語群の尾高型出現率、③全体の尾高型出現率、④較差(③-②)は、次の通りである。括弧内は該当語数を示す。

	《尾高4》	《尾高3》	《尾高2》	《尾高1》	《尾高0》
①	66.1 (9)	67.2 (26)	71.4 (7)	68.4 (9)	54.6 (8)
②	14.0 (93)	19.3 (74)	10.4 (46)	8.4 (75)	12.3 (128)
③	18.6 (102)	31.7 (100)	18.5 (53)	14.8 (84)	14.8 (136)
④	4.6	12.4	8.1	6.4	2.5

- ①から、「～・書」「～・所」「～・署」を後部要素とする複合語は、《尾高3》に限らず、全体的に尾高型出現率がきわめて高いことが分かる。一方、②から明らかなように、全体から①を除いたグループの出現率は、いずれも20%以下にすぎない。②に①を加えることによって③の全体となるが、その際に上昇幅が最も大きいのが《尾高3》であることを、④の数値が示している。《尾高3》が特に目立つが、それ以外でもそれなりの数値の上昇がみられることを補足しておく。
- 4 辞書のアクセントでも、『新明解』のような国語辞典と、『NHK』『明解ア』『全国ア』のようなアクセント辞典とは、さらに立項基準が異なるはずである。表1の4種の辞書における「x」の分布からも、その一端がうかがえよう。

# 参考文献

- 相澤 正夫 (1991) 「生きているアクセント規則の検討—東京語の単純動詞とその転成名詞の場合—」『研究報告集 12』(国立国語研究所報告 103)
- (1992) 「進行中のアクセント変化—東京語の複合動詞の場合—」『研究報告集 13』(国立国語研究所報告 104)
- (1996a) 「語の長さアクセント変化—『東京語アクセント資料』の分析—」『国立国語研究所研究報告集 17』
- (1996b) 「尾高型アクセントの現在位置—『東京語アクセント資料』の分析—」『言語学林 1995-1996』(柴田武先生喜寿記念論文集) 三省堂
- 佐藤 亮一 (1990) 「現代東京語のアクセント—年齢差および辞典との差を中心に—」『国語論究 2 文字・音韻の研究』明治書院
- 杉藤 美代子・田原 広史 (1989) 「統計的観点から見た大阪アクセント—東京との比較を中心に—」『音声言語 3』(近畿音声言語研究会)
- 馬瀬 良雄・佐藤 亮一 (1989) 「東京語アクセントの多様性」『講座日本語と日本語教育 2 日本語の音声・音韻 (上)』明治書院

資料集・辞典類（見出しは本稿での略称）

- 『東京ア』：『東京語アクセント資料 上・下』（文部省科学研究費特定研究「言語の標準化」資料集）柴田武監修，馬瀬良雄・佐藤亮一編（1985）
- 『新明解』：『新明解国語辞典 第3版』（第1刷）金田一京助ほか編（1981）三省堂
- 『NHK』：『日本語発音アクセント辞典』（第15刷）日本放送協会編（1974）日本放送出版協会
- 『明解ア』：『明解日本語アクセント辞典 第2版』（第1刷）金田一春彦監修，秋永一枝編（1981）三省堂
- 『全国ア』：『全国アクセント辞典』（第20版）平山輝男編（1979）東京堂出版

（原稿受理日：1997年1月20日）

---

相澤 正夫（あいざわ まさお）

国立国語研究所日本語教育センター 115 東京都北区西が丘3-9-14  
aiz@kokken. go. jp

## ***A Dictionary of Tone-accent on Words in the Tokyo Dialect* and accent information in four dictionaries:**

In search of a methodology for data evaluation

AIZAWA Masao

The National Language Research Institute

### **Key words**

accent in the Tokyo dialect, word final accent, accent information in dictionaries,  
data evaluation, *A Dictionary of Tone-accent on Words in the Tokyo Dialect* I, II

*A Dictionary of Tone-accent on Words in the Tokyo Dialect* (1985) presents in tabular form a large amount of accent data with 12,803 word entries. Each of the entries lists the accent patterns obtained from each of nineteen informants who vary in age, sex, and the area in which they were raised (*yamanote* 'uptown' or *shitamachi* 'downtown'), as well as the accent patterns registered in four widely available dictionaries currently in use.

The accent information in the four dictionaries mentioned above is very useful for the evaluation of accent data contained in this new dictionary in that we can easily examine the differences among them by comparing them to one another quantitatively or statistically.

In this paper, making the most of this wealth of data, I seek a methodology for evaluating accent data. Taking the appearance of the word-final accent in the accent pattern system of nouns as a case study, I examine the data exhaustively and discuss the following points:

1. As a whole, the accent information in the new dictionary corresponds well to that in the four commercial dictionaries; therefore we can make good use of the former as the most up-to-date accent data for comparison with the latter.
2. However, the selection of the word entries in the new dictionary is idiosyncratic; therefore we should be careful when we use it as a source for statistical studies.
3. To give an example, the new dictionary contains many compound nouns which have a very productive suffixing element in common. As a result, many of the compound nouns with the same accent pattern are produced through the application of the same accentuation rule, causing a statistical skewing in the data.